Chapter 9

The Spectrum of Bounded Linear Operators

In Chapter 7, we used Fourier series to solve various constant coefficient, linear partial differential equations, such as the heat equation. Consider, as an example, the following initial boundary value problem for a variable coefficient, linear equation

$$\begin{aligned} &u_t = u_{xx} - q(x)u && 0 < x < 1,\, t > 0, \\ &u(0,t) = 0, \quad u(1,t) = 0 && t \geq 0, \\ &u(x,0) = f(x) && 0 \leq x \leq 1, \end{aligned}$$

where q is a given coefficient function. This equation describes the temperature of a heat conducting bar with a nonuniform heat loss term given by $-q(x)u$. What would it take to express the solution for given initial data f as a series expansion similar to a Fourier series?

If we use separation of variables and look for a solution of the form

$$u(x,t) = \sum_{n=1}^{\infty} a_n(t) u_n(x),$$

where $\{u_n \mid n \in \mathbb{N}\}$ is a basis of $L^2([0,1])$, then we find that the a_n satisfy

$$\frac{da_n}{dt} = -\lambda_n a_n$$

for some constants λ_n, and the u_n satisfy

$$-\frac{d^2 u_n}{dx^2} + q u_n = \lambda_n u_n.$$

Thus, the u_n should be eigenvectors of the linear operator A defined by

$$\begin{aligned} Au &= -\frac{d^2 u}{dx^2} + qu, \\ u(0) &= 0, \quad u(1) = 0. \end{aligned}$$

We therefore want to find a complete set of eigenvectors of A, or, equivalently, to diagonalize A. The problem of diagonalizing a linear map on an infinite-dimensional space arises in many other ways, and is part of what is called spectral theory.

Spectral theory provides a powerful way to understand linear operators by decomposing the space on which they act into invariant subspaces on which their action is simple. In the finite-dimensional case, the spectrum of a linear operator consists of its eigenvalues. The action of the operator on the subspace of eigenvectors with a given eigenvalue is just multiplication by the eigenvalue. As we will see, the spectral theory of bounded linear operators on infinite-dimensional spaces is more involved. For example, an operator may have a *continuous spectrum* in addition to, or instead of, a *point spectrum* of eigenvalues. A particularly simple and important case is that of compact, self-adjoint operators. Compact operators may be approximated by finite-dimensional operators, and their spectral theory is close to that of finite-dimensional operators. We begin with a brief review of the finite-dimensional case.

9.1 Diagonalization of matrices

We consider an $n \times n$ matrix A with complex entries as a bounded linear map $A : \mathbb{C}^n \to \mathbb{C}^n$. A complex number λ is an *eigenvalue* of A if there is a nonzero vector $u \in \mathbb{C}^n$ such that

$$Au = \lambda u. \tag{9.1}$$

A vector $u \in \mathbb{C}^n$ such that (9.1) holds is called an *eigenvector* of A associated with the eigenvalue λ.

The matrix A is *diagonalizable* if there is a basis $\{u_1, \ldots, u_n\}$ of $\mathbb{C}^n$ consisting of eigenvectors of A, meaning that there are eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$ in $\mathbb{C}$, which need not be distinct, such that

$$Au_k = \lambda_k u_k \qquad \text{for } k = 1, \ldots, n. \tag{9.2}$$

The set of eigenvalues of A is called the *spectrum* of A, and is denoted by $\sigma(A)$. The most useful bases of Hilbert spaces are orthonormal bases. A natural question is therefore: When does an $n \times n$ matrix have an orthonormal basis of eigenvectors?

If $\{u_1, \ldots, u_n\}$ is an orthonormal basis of $\mathbb{C}^n$, then the matrix $U = (u_1, \ldots, u_n)$, whose columns are the basis vectors, is a unitary matrix such that

$$Ue_k = u_k, \qquad U^* u_k = e_k,$$

where $\{e_1, \ldots, e_n\}$ is the standard basis of $\mathbb{C}^n$. If the basis vectors $\{u_1, \ldots, u_n\}$ are eigenvectors of A, as in (9.2), then

$$U^* A U e_k = \lambda_k e_k.$$

It follows that U^*AU is a diagonal matrix with the eigenvalues of A on the diagonal, so $A = UDU^*$ where

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Conversely, if $A = UDU^*$ with U unitary and D diagonal, then the columns of U form an orthonormal basis of $\mathbb{C}^n$ consisting of eigenvectors of A. Thus, a matrix A can be diagonalized by a unitary matrix if and only if $\mathbb{C}^n$ has an orthonormal basis consisting of eigenvectors of A.

If $A = UDU^*$, then $A^* = U\overline{D}U^*$. Since any two diagonal matrices commute, it follows that A commutes with its Hermitian conjugate A^*:

$$A^*A = U\overline{D}U^*UDU^* = U\overline{D}DU^* = UD\overline{D}U^* = UDU^*U\overline{D}U^* = AA^*.$$

Matrices with this property are called *normal matrices*. For example, Hermitian matrices A, satisfying $A^* = A$, and unitary matrices U, satisfying $U^*U = I$, are normal. We have shown that any matrix with an orthonormal basis of eigenvectors is normal. A standard theorem in linear algebra, which we will not prove here, is that the converse also holds.

Theorem 9.1 An $n \times n$ complex matrix A is normal if and only if $\mathbb{C}^n$ has an orthonormal basis consisting of eigenvectors of A.

A normal matrix N can be written as the product of a unitary matrix V, and a nonnegative, Hermitian matrix A. This follows directly from the diagonal form of N. If $N = UDU^*$ has eigenvalues $\lambda_k = |\lambda_k|e^{i\varphi_k}$, we can write $D = \Phi|D|$, where Φ is a diagonal matrix with entries $e^{i\varphi_k}$ and $|D|$ a diagonal matrix with entries $|\lambda_k|$. Then

$$N = VA, \tag{9.3}$$

where $V = U\Phi U^*$ is unitary, and $A = U|D|U^*$ is nonnegative, meaning that,

$$u^*Au \geq 0 \qquad \text{for all } u \in \mathbb{C}^n.$$

It is straightforward to check that $VA = AV$. Equation (9.3) is called the *polar decomposition* of N. It is a matrix analog of the polar decomposition of a complex number $z = re^{i\theta}$ as the product of a nonnegative number r and a complex number $e^{i\theta}$ on the unit circle. The converse is also true: if $N = VA$, with V unitary, A Hermitian, and $VA = AV$, then N is normal.

The eigenvalues of a matrix A are the roots of the *characteristic polynomial* p_A of A, given by

$$p_A(\lambda) = \det(A - \lambda I).$$

If $p_A(\lambda) = 0$, then $A - \lambda I$ is singular, so $\ker(A - \lambda I) \neq \{0\}$ and there is an associated eigenvector. Since every polynomial has at least one root, it follows that every matrix has at least one eigenvalue, and each distinct eigenvalue has a nonzero subspace of eigenvectors.

It is not true that all matrices have a basis of eigenvectors, because the dimension of the space of eigenvectors associated with a multiple root of the characteristic polynomial may be strictly less than the algebraic multiplicity of the root. We call the dimension of the eigenspace associated with a given eigenvalue the *geometric multiplicity* of the eigenvalue, or the *multiplicity* for short.

Example 9.2 The 2×2 Jordan block

$$A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

has one eigenvalue λ. The eigenvectors associated with λ are scalar multiples of $u = (1, 0)^T$, so its multiplicity is one, and the eigenspace does not include a basis of $\mathbb{C}^2$. The matrix is not normal, since

$$[A, A^*] = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

9.2 The spectrum

A bounded linear operator on an infinite-dimensional Hilbert space need not have any eigenvalues at all, even if it is self-adjoint (see Example 9.5 below). Thus, we cannot hope to find, in general, an orthonormal basis of the space consisting entirely of eigenvectors. It is therefore necessary to define the spectrum of a linear operator on an infinite-dimensional space in a more general way than as the set of eigenvalues. We denote the space of bounded linear operators on a Hilbert space $\mathcal{H}$ by $\mathcal{B}(\mathcal{H})$.

Definition 9.3 The *resolvent set* of an operator $A \in \mathcal{B}(\mathcal{H})$, denoted by $\rho(A)$, is the set of complex numbers λ such that $(A - \lambda I) : \mathcal{H} \to \mathcal{H}$ is one-to-one and onto. The *spectrum* of A, denoted by $\sigma(A)$, is the complement of the resolvent set in $\mathbb{C}$, meaning that $\sigma(A) = \mathbb{C} \setminus \rho(A)$.

If $A - \lambda I$ is one-to-one and onto, then the open mapping theorem implies that $(A - \lambda I)^{-1}$ is bounded. Hence, when $\lambda \in \rho(A)$, both $A - \lambda I$ and $(A - \lambda I)^{-1}$ are one-to-one, onto, bounded linear operators.

As in the finite-dimensional case, a complex number λ is called an eigenvalue of A if there is a nonzero vector $u \in \mathcal{H}$ such that $Au = \lambda u$. In that case, $\ker(A - \lambda I) \neq \{0\}$, so $A - \lambda I$ is not one-to-one, and $\lambda \in \sigma(A)$. This is not the only way, however, that a complex number can belong to the spectrum. We subdivide the spectrum of a bounded linear operator as follows.

Definition 9.4 Suppose that A is a bounded linear operator on a Hilbert space $\mathcal{H}$.

(a) The *point spectrum* of A consists of all $\lambda \in \sigma(A)$ such that $A - \lambda I$ is not one-to-one. In this case λ is called an *eigenvalue* of A.
(b) The *continuous spectrum* of A consists of all $\lambda \in \sigma(A)$ such that $A - \lambda I$ is one-to-one but not onto, and $\operatorname{ran}(A - \lambda I)$ is dense in $\mathcal{H}$.
(c) The *residual spectrum* of A consists of all $\lambda \in \sigma(A)$ such that $A - \lambda I$ is one-to-one but not onto, and $\operatorname{ran}(A - \lambda I)$ is not dense in $\mathcal{H}$.

The following example gives a bounded, self-adjoint operator whose spectrum is purely continuous.

Example 9.5 Let $\mathcal{H} = L^2([0,1])$, and define the multiplication operator $M : \mathcal{H} \to \mathcal{H}$ by

$$Mf(x) = xf(x).$$

Then M is bounded with $\|M\| = 1$. If $Mf = \lambda f$, then $f(x) = 0$ a.e., so $f = 0$ in $L^2([0,1])$. Thus, M has no eigenvalues. If $\lambda \notin [0,1]$, then $(x-\lambda)^{-1}f(x) \in L^2([0,1])$ for any $f \in L^2([0,1])$ because $(x - \lambda)$ is bounded away from zero on $[0,1]$. Thus, $\mathbb{C} \setminus [0,1]$ is in the resolvent set of M. If $\lambda \in [0,1]$, then $M - \lambda I$ is not onto, because $c(x-\lambda)^{-1} \notin L^2([0,1])$ for $c \neq 0$, so the nonzero constant functions c do not belong to the range of $M - \lambda I$. The range of $M - \lambda I$ is dense, however. For any $f \in L^2([0,1])$, let

$$f_n(x) = \begin{cases} f(x) & \text{if } |x - \lambda| \geq 1/n, \\ 0 & \text{if } |x - \lambda| < 1/n. \end{cases}$$

Then $f_n \to f$ in $L^2([0,1])$, and $f_n \in \operatorname{ran}(M - \lambda I)$, since $(x-\lambda)^{-1}f_n(x) \in L^2([0,1])$. It follows that $\sigma(M) = [0,1]$, and that every $\lambda \in [0,1]$ belongs to the continuous spectrum of M. If M acts on the "delta function" supported at λ, which is a distribution (see Chapter 11) with the property that for every continuous function f,

$$f(x)\delta_\lambda(x) = f(\lambda)\delta_\lambda(x),$$

then $M\delta_\lambda = \lambda\delta_\lambda$. Thus, in some sense, there are eigenvectors associated with points in the continuous spectrum of M, but they lie outside the space $L^2([0,1])$ on which M acts.

If λ belongs to the resolvent set $\rho(A)$ of a linear operator A, then $A - \lambda I$ has an everywhere defined, bounded inverse. The operator

$$R_\lambda = (\lambda I - A)^{-1} \tag{9.4}$$

is called the *resolvent* of A at λ, or simply the resolvent of A. The resolvent of A is an operator-valued function defined on the subset $\rho(A)$ of $\mathbb{C}$.

An operator-valued function $F : \Omega \to \mathcal{B}(\mathcal{H})$, defined on an open subset Ω of the complex plane $\mathbb{C}$, is said to be *analytic* at $z_0 \in \Omega$ if there are operators $F_n \in \mathcal{B}(\mathcal{H})$ and a $\delta > 0$ such that

$$F(z) = \sum_{n=0}^{\infty} (z - z_0)^n F_n,$$

where the power series on the right-hand side converges with respect to the operator norm on $\mathcal{B}(\mathcal{H})$ in a disc $|z - z_0| < \delta$ for some $\delta > 0$. We say that F is *analytic* or *holomorphic* in Ω if it is analytic at every point in Ω. This definition is a straightforward generalization of the definition of an analytic complex-valued function $f : \Omega \subset \mathbb{C} \to \mathbb{C}$ as a function with a convergent power series expansion at each point of Ω. The fact that we are dealing with vector-valued, or operator-valued, functions instead of complex-valued functions makes very little difference.

Proposition 9.6 If A is a bounded linear operator on a Hilbert space, then the resolvent set $\rho(A)$ is an open subset of $\mathbb{C}$ that contains the exterior disc $\{\lambda \in \mathbb{C} \mid |\lambda| > \|A\|\}$. The resolvent R_λ is an operator-valued analytic function of λ defined on $\rho(A)$.

Proof. Suppose that $\lambda_0 \in \rho(A)$. Then we may write

$$\lambda I - A = (\lambda_0 I - A)\left[I - (\lambda_0 - \lambda)(\lambda_0 I - A)^{-1}\right].$$

If $|\lambda_0 - \lambda| < \|(\lambda_0 I - A)^{-1}\|^{-1}$, then we can invert the operator on the right-hand side by the Neumann series (see Exercise 5.17). Hence, there is an open disk in the complex plane with center λ_0 that is contained in $\rho(A)$. Moreover, the resolvent R_λ is given by an operator-norm convergent Taylor series in the disc, so it is analytic in $\rho(A)$. If $|\lambda| > \|A\|$, then the Neumann series also shows that $R_\lambda = \lambda\,(I - A/\lambda)$ is invertible, so $\lambda \in \rho(A)$. □

Since the spectrum $\sigma(A)$ of A is the complement of the resolvent set, it follows that the spectrum is a closed subset of $\mathbb{C}$, and

$$\sigma(A) \subset \{z \in \mathbb{C} \mid |z| \leq \|A\|\}.$$

The *spectral radius* of A, denoted by $r(A)$, is the radius of the smallest disk centered at zero that contains $\sigma(A)$,

$$r(A) = \sup\{|\lambda| \mid \lambda \in \sigma(A)\}.$$

We can refine Proposition 9.6 as follows.

Proposition 9.7 If A is a bounded linear operator, then

$$r(A) = \lim_{n\to\infty} \|A^n\|^{1/n}. \tag{9.5}$$

If A is self-adjoint, then $r(A) = \|A\|$.

Proof. To prove (9.5), we first show that the limit on the right-hand side exists. Let

$$a_n = \log \|A^n\|.$$

We want to show that (a_n/n) converges. Since $\|A^{m+n}\| \leq \|A^m\| \, \|A^n\|$, we have $a_n \leq na_1$ and

$$a_{m+n} \leq a_m + a_n.$$

We write $n = pm + q$ where $0 \leq q < m$. It follows that

$$\frac{a_n}{n} \leq \frac{p}{n} a_m + \frac{1}{n} a_q.$$

If $n \to \infty$ with m fixed, then $p/n \to 1/m$, so

$$\limsup_{n\to\infty} \frac{a_n}{n} \leq \frac{a_m}{m}.$$

Taking the limit as $m \to \infty$, we obtain that

$$\limsup_{n\to\infty} \frac{a_n}{n} \leq \liminf_{m\to\infty} \frac{a_m}{m},$$

which implies that (a_n/n) converges.

Equation (9.5) implies that the Neumann series

$$I + A + A^2 + \ldots + A^n + \ldots$$

converges if $r(A) < 1$ and diverges if $r(A) > 1$: if $r(A) < 1$, then there is an $r(A) < R < 1$ and an N such that $\|A^n\| \leq R^n$ for all $n \geq N$; while if $r(A) > 1$, there is an $1 < R < r(A)$ and an N such that $\|A^n\| \geq R^n$ for all $n \geq N$. It follows that $\lambda I - A$ may be inverted by a Neumann series when $|\lambda| > r(A)$, so the spectrum of A is contained inside the disc $\{\lambda \in \mathbb{C} \mid |\lambda| \leq r(A)\}$, and that the Neumann series must diverge, so $\lambda I - A$ is not invertible, for some $\lambda \in \mathbb{C}$ with $|\lambda| = r(A)$. For more details, see Reed and Simon [45], for example.

From Corollary 8.27, if A is self-adjoint, then $\|A^2\| = \|A\|^2$. The repeated use of this result implies that $\|A^{2^m}\| = \|A\|^{2^m}$, and hence (9.5), applied to the subsequence with $n = 2^m$, implies that $r(A) = \|A\|$. □

Although the spectral radius of a self-adjoint operator is equal to its norm, the spectral radius does not provide a norm on the space of all bounded operators. In particular, $r(A) = 0$ does not imply that $A = 0$, as Exercise 5.13 illustrates. If $r(A) = 0$, then A is called a *nilpotent operator*.

Proposition 9.8 The spectrum of a bounded operator on a Hilbert space is nonempty.

Proof. Suppose that $A \in \mathcal{B}(\mathcal{H})$. Then the resolvent $R_\lambda = (\lambda I - A)^{-1}$ is an analytic function on $\rho(A)$. Therefore, for every $x, y \in \mathcal{H}$, the function $f : \rho(A) \to \mathbb{C}$ defined by

$$f(\lambda) = \langle x, R_\lambda y \rangle$$

is analytic in $\rho(A)$, and $\lim_{\lambda \to \infty} f(\lambda) = 0$. Suppose, for contradiction, that $\sigma(A)$ is empty. Then f is a bounded entire function, and Liouville's Theorem implies that $f : \mathbb{C} \to \mathbb{C}$ is a constant function, so $f = 0$. But if $f = 0$ for every $x, y \in \mathcal{H}$, then $R_\lambda = 0$ for all $\lambda \in \mathbb{C}$, which is impossible. Hence, $\sigma(A)$ is not empty. □

The spectrum of a bounded operator may, however, consist of a single point (see Exercise 9.7 for an example).

9.3 The spectral theorem for compact, self-adjoint operators

In this section, we analyze the spectrum of a compact, self-adjoint operator. The spectrum consists entirely of eigenvalues, with the possible exception of zero, which may belong to the continuous spectrum. We begin by proving some basic properties of the spectrum of a bounded, self-adjoint operator.

Lemma 9.9 The eigenvalues of a bounded, self-adjoint operator are real, and eigenvectors associated with different eigenvalues are orthogonal.

Proof. If $A : \mathcal{H} \to \mathcal{H}$ is self-adjoint, and $Ax = \lambda x$ with $x \neq 0$, then

$$\lambda \langle x, x \rangle = \langle x, Ax \rangle = \langle Ax, x \rangle = \overline{\lambda} \langle x, x \rangle,$$

so $\lambda = \overline{\lambda}$, and $\lambda \in \mathbb{R}$. If $Ax = \lambda x$ and $Ay = \mu y$, where λ and μ are real, then

$$\lambda \langle x, y \rangle = \langle Ax, y \rangle = \langle x, Ay \rangle = \mu \langle x, y \rangle.$$

It follows that if $\lambda \neq \mu$, then $\langle x, y \rangle = 0$ and $x \perp y$. □

As we will see in the next chapter, self-adjoint operators are a rich source of orthonormal bases.

A linear subspace $\mathcal{M}$ of $\mathcal{H}$ is called an *invariant subspace* of a linear operator A on $\mathcal{H}$ if $Ax \in \mathcal{M}$ for all $x \in \mathcal{M}$. In that case, the restriction $A|_{\mathcal{M}}$ of A to $\mathcal{M}$ is a linear operator on $\mathcal{M}$. Suppose that $\mathcal{H} = \mathcal{M} \oplus \mathcal{N}$ is a direct sum of invariant subspaces $\mathcal{M}$ and $\mathcal{N}$ of A. Then every $x \in \mathcal{H}$ may be written as $x = y + z$, with $y \in \mathcal{M}$ and $z \in \mathcal{N}$, and

$$Ax = A|_{\mathcal{M}} y + A|_{\mathcal{N}} z. \tag{9.6}$$

Thus, the action of A on $\mathcal{H}$ is determined by its actions on the invariant subspaces.

Example 9.10 Consider matrices acting on $\mathbb{C}^d = \mathbb{C}^m \oplus \mathbb{C}^n$ where $d = m + n$. A $d \times d$ matrix A leaves $\mathbb{C}^m$ invariant if it has the block form

$$A = \begin{pmatrix} B & D \\ 0 & C \end{pmatrix},$$

where B is an $m \times m$ matrix, D is $m \times n$, and C is $n \times n$. The matrix A leaves both $\mathbb{C}^m$ and the complementary space $\mathbb{C}^n$ invariant if $D = 0$.

An invariant subspace of a nondiagonalizable operator may have no complementary invariant subspace. However, the orthogonal complement of an invariant subspace of a self-adjoint operator is also invariant, as we prove in the following lemma. Thus, we can decompose the action of a self-adjoint operator on a linear space into actions on smaller orthogonal invariant subspaces.

Lemma 9.11 If A is a bounded, self-adjoint operator on a Hilbert space $\mathcal{H}$ and $\mathcal{M}$ is an invariant subspace of A, then $\mathcal{M}^\perp$ is an invariant subspace of A.

Proof. If $x \in \mathcal{M}^\perp$ and $y \in \mathcal{M}$, then

$$\langle y, Ax \rangle = \langle Ay, x \rangle = 0$$

because $A = A^*$ and $Ay \in \mathcal{M}$. Therefore, $Ax \in \mathcal{M}^\perp$. □

Next, we show that the whole spectrum — not just the point spectrum — of a bounded, self-adjoint operator is real, and that the residual spectrum is empty. We begin with a preliminary proposition.

Proposition 9.12 If λ belongs to the residual spectrum of a bounded operator A on a Hilbert space, then $\overline{\lambda}$ is an eigenvalue of A^*.

Proof. If λ belongs to the residual spectrum of $A \in \mathcal{B}(\mathcal{H})$, then $\operatorname{ran}(A - \lambda I)$ is not dense in $\mathcal{H}$. By Theorem 6.13, there is a nonzero vector $x \in \mathcal{H}$ such that $x \perp \operatorname{ran}(A - \lambda I)$. Theorem 8.17 then implies that $x \in \ker\left(A^* - \overline{\lambda} I\right)$. □

Lemma 9.13 If A is a bounded, self-adjoint operator on a Hilbert space, then the spectrum of A is real and is contained in the interval $[-\|A\|, \|A\|]$.

Proof. We have shown that $r(A) \leq \|A\|$, so we only have to prove that the spectrum is real. Suppose that $\lambda = a + ib \in \mathbb{C}$, where with $a, b \in \mathbb{R}$ and $b \neq 0$. For any $x \in \mathcal{H}$, we have

$$\begin{aligned}
\|(A - \lambda I)x\|^2 &= \langle (A - \lambda I)x, (A - \lambda I)x \rangle \\
&= \langle (A - aI)x, (A - aI)x \rangle + \langle (-ib)x, (-ib)x \rangle \\
&\quad + \langle Ax, (-ib)x \rangle + \langle (-ib)x, Ax \rangle \\
&= \|(A - aI)x\|^2 + b^2 \|x\|^2 \\
&\geq b^2 \|x\|^2.
\end{aligned}$$

It follows from this estimate and Proposition 5.30 that $A - \lambda I$ is one-to-one and has closed range. If $\operatorname{ran}(A - \lambda I) \neq \mathcal{H}$, then λ belongs to the residual spectrum of A, and, by Proposition 9.12, $\overline{\lambda} = a - ib$ is an eigenvalue of A. Thus A has an eigenvalue that does not belong to $\mathbb{R}$, which contradicts Lemma 9.9. It follows that $\lambda \in \rho(A)$ if λ is not real. □

Corollary 9.14 The residual spectrum of a bounded, self-adjoint operator is empty.

Proof. From Lemma 9.13, the point spectrum and the residual spectrum are disjoint subsets of $\mathbb{R}$, so the result follows immediately from Proposition 9.12. □

Bounded linear operators on an infinite-dimensional Hilbert space do not always behave like operators on a finite-dimensional space. We have seen in Example 9.5 that a bounded, self-adjoint operator may have no eigenvalues, while the identity operator on an infinite-dimensional Hilbert space has a nonzero eigenvalue of infinite multiplicity. The properties of compact operators are much closer to those of operators on finite-dimensional spaces, and we will study their spectral theory next.

Proposition 9.15 A nonzero eigenvalue of a compact, self-adjoint operator has finite multiplicity. A countably infinite set of nonzero eigenvalues has zero as an accumulation point, and no other accumulation points.

Proof. Suppose, for contradiction, that λ is a nonzero eigenvalue with infinite multiplicity. Then there is a sequence (e_n) of orthonormal eigenvectors. This sequence is bounded, but (Ae_n) does not have a convergent subsequence because $Ae_n = \lambda e_n$, which contradicts the compactness of A.

If A has a countably infinite set $\{\lambda_n\}$ of nonzero eigenvalues, then, since the eigenvalues are bounded by $\|A\|$, there is a convergent subsequence (λ_{n_k}). If $\lambda_{n_k} \to \lambda$ and $\lambda \neq 0$, then the orthogonal sequence of eigenvectors (f_{n_k}), where $f_{n_k} = \lambda_{n_k}^{-1} e_{n_k}$ and $\|e_{n_k}\| = 1$, would be bounded; but (Af_{n_k}) has no convergent subsequence since $Af_{n_k} = e_{n_k}$. □

To motivate the statement of the spectral theorem for compact, self-adjoint operators, suppose that $x \in \mathcal{H}$ is given by

$$x = \sum_k c_k e_k + z, \tag{9.7}$$

where $\{e_k\}$ is an orthonormal set of eigenvectors of A with corresponding nonzero eigenvalues $\{\lambda_k\}$, $z \in \ker A$, and $c_k \in \mathbb{C}$. Then $Ax = \sum_k \lambda_k c_k e_k$. Let P_k denote the one-dimensional orthogonal projection onto the subspace spanned by e_k,

$$P_k x = \langle e_k, x\rangle e_k. \tag{9.8}$$

From Lemma 9.9, we have $z \perp e_k$, so $c_k = \langle e_k, x\rangle$ and

$$Ax = \sum_k \lambda_k P_k x. \tag{9.9}$$

If λ_k has finite multiplicity $m_k > 1$, meaning that the dimension of the associated eigenspace $E_k \subset \mathcal{H}$ is greater than one, then we may combine the one-dimensional projections associated with the same eigenvalues. In doing so, we may represent A by a sum of the same form as (9.9) in which the λ_k are distinct, and the P_k are orthogonal projections onto the eigenspaces E_k.

The spectral theorem for compact, self-adjoint operators states that any $x \in \mathcal{H}$ can be expanded in the form (9.7), and that A can be expressed as a sum of orthogonal projections, as in (9.9).

Theorem 9.16 (Spectral theorem for compact, self-adjoint operators) Let $A : \mathcal{H} \to \mathcal{H}$ be a compact, self-adjoint operator on a Hilbert space $\mathcal{H}$. There is an orthonormal basis of $\mathcal{H}$ consisting of eigenvectors of A. The nonzero eigenvalues of A form a finite or countably infinite set $\{\lambda_k\}$ of real numbers, and

$$A = \sum_k \lambda_k P_k, \tag{9.10}$$

where P_k is the orthogonal projection onto the finite-dimensional eigenspace of eigenvectors with eigenvalue λ_k. If the number of nonzero eigenvalues is countably infinite, then the series in (9.10) converges to A in the operator norm.

Proof. First we prove that if A is compact and self-adjoint, then $\lambda = \|A\|$ or $\lambda = -\|A\|$ (or both) is an eigenvalue of A. This is the crucial part of the proof. We use a variational argument to obtain an eigenvector.

There is nothing to prove if $A = 0$, so we suppose that $A \neq 0$. From Lemma 8.26, we have

$$\|A\| = \sup_{\|x\|=1} |\langle x, Ax\rangle|.$$

Hence, there is a sequence (x_n) in $\mathcal{H}$ with $\|x_n\| = 1$ such that

$$\|A\| = \lim_{n\to\infty} |\langle x_n, Ax_n\rangle|.$$

Since A is self-adjoint, $\langle x_n, Ax_n\rangle$ is real for all n, so there is a subsequence of (x_n), which we still denote by (x_n), such that

$$\lim_{n\to\infty} \langle x_n, Ax_n\rangle = \lambda, \tag{9.11}$$

where $\lambda = \|A\|$ or $\lambda = -\|A\|$.

The sequence (x_n) consists of unit vectors, so it is bounded. The compactness of A implies that there is a subsequence, which we still denote by (x_n), such that (Ax_n) converges. We let $y = \lim_{n\to\infty} Ax_n$. We claim that y is an eigenvector of A with eigenvalue λ. First, $y \neq 0$, since otherwise (9.11) would imply that $\lambda = 0$, which is not the case since $|\lambda| = \|A\|$ and $A \neq 0$. The fact that y is an eigenvector follows from the following computation:

$$\begin{aligned}
\|(A-\lambda I)y\|^2 &= \lim_{n\to\infty} \|(A-\lambda I)Ax_n\|^2 \\
&\leq \|A\|^2 \lim_{n\to\infty} \|(A-\lambda I)x_n\|^2 \\
&= \|A\|^2 \lim_{n\to\infty} \left[\|Ax_n\|^2 + \lambda^2\|x_n\|^2 - 2\lambda\langle x_n, Ax_n\rangle\right] \\
&\leq \|A\|^2 \lim_{n\to\infty} \left[\|A\|^2\|x_n\|^2 + \lambda^2\|x_n\|^2 - 2\lambda\langle x_n, Ax_n\rangle\right] \\
&= \|A\|^2 \left[\lambda^2 + \lambda^2 - 2\lambda^2\right] \\
&= 0.
\end{aligned}$$

To complete the proof, we use an orthogonal decomposition of $\mathcal{H}$ into invariant subspaces to apply the result we have just proved to smaller and smaller subspaces of $\mathcal{H}$. We let $\mathcal{N}_1 = \mathcal{H}$ and $A_1 = A$. There is a normalized eigenvector of A_1, which we denote by e_1, with eigenvalue λ_1, where $|\lambda_1| = \|A_1\|$. Let $\mathcal{M}_2$ be the one-dimensional subspace of $\mathcal{N}_1$ generated by e_1. Then $\mathcal{M}_2$ is an invariant subspace of A_1. We decompose $\mathcal{N}_1 = \mathcal{M}_2 \oplus \mathcal{N}_2$, where $\mathcal{N}_2 = \mathcal{M}_2^\perp$. Lemma 9.11 implies that $\mathcal{N}_2$ is an invariant subspace of A_1. We denote the restriction of A_1 to $\mathcal{N}_2$ by A_2. Then A_2 is the difference of two compact operators, so it is compact by Proposition 5.43. We also have that $\|A_2\| \leq \|A_1\|$, since $\mathcal{N}_2 \subset \mathcal{N}_1$.

An application of the same argument to A_2 implies that A_2 has a normalized eigenvector e_2 with eigenvalue λ_2, where

$$|\lambda_2| = \|A_2\| \leq \|A_1\| = |\lambda_1|.$$

Moreover, $e_2 \perp e_1$. Repeating this procedure, we define A_n inductively to be the restriction of A to $\mathcal{N}_n = \mathcal{M}_n^\perp$, where $\mathcal{M}_n$ is the space spanned by $\{e_1, \ldots e_{n-1}\}$, and e_n to be an eigenvector of A_n with eigenvalue λ_n. By construction, $|\lambda_n| = \|A_n\|$ and $(|\lambda_n|)$ is a nonincreasing sequence.

If $A_{n+1} = 0$ for some n, then A has only finitely many nonzero eigenvalues, and it is given by a finite sum of the form (9.10). If $\dim \mathcal{H} > n$, then the orthonormal set $\{e_k \mid k = 1, \ldots n\}$ can be extended to an orthonormal basis of $\mathcal{H}$. All other basis vectors are eigenvectors of A with eigenvalue zero.

If $A_n \neq 0$ for every $n \in \mathbb{N}$, then we obtain an infinite sequence of nonzero eigenvalues and eigenvectors. From Proposition 9.15, the eigenvalues have finite multiplicities and

$$\lim_{n\to\infty} \lambda_n = 0. \tag{9.12}$$

For any $n \in \mathbb{N}$, we have

$$A = \sum_{k=1}^{n} \lambda_k P_k + A_{n+1},$$

where A_{n+1} is zero on the subspace spanned by $\{e_1, \ldots, e_n\}$, and $\|A_{n+1}\| = |\lambda_{n+1}|$. By (9.12), we have

$$\lim_{n\to\infty} \left\| A - \sum_{k=1}^{n} \lambda_k \langle e_k, \cdot \rangle e_k \right\| = \lim_{n\to\infty} |\lambda_{n+1}| = 0,$$

meaning that

$$A = \sum_{k=1}^{\infty} \lambda_k P_k,$$

where the sum converges in the operator norm.

If A has an infinite sequence of nonzero eigenvalues, then the range of A is

$$\operatorname{ran} A = \left\{ \sum_{k=1}^{\infty} c_k e_k \;\middle|\; \sum_{k=1}^{\infty} \frac{|c_k|^2}{|\lambda_k|^2} < \infty \right\}.$$

The range is not closed since $\lambda_n \to 0$ as $n \to \infty$. The closure of the range, $\mathcal{M} = \overline{\operatorname{ran} A}$, is the closed linear span of the set of eigenvectors $\{e_n \mid n \in \mathbb{N}\}$ with nonzero eigenvalues,

$$\mathcal{M} = \left\{ \sum_{k=1}^{\infty} c_k e_k \;\middle|\; \sum_{k=1}^{\infty} |c_k|^2 < \infty \right\}.$$

If $x \in \mathcal{M}^\perp$, then $Ax = A_n x$ for all $n \in \mathbb{N}$, so that

$$\|Ax\| \le \|A_n\| \, \|x\| \to 0 \qquad \text{as } n \to \infty.$$

Therefore, $\mathcal{M}^\perp = \ker A$ consists of eigenvectors of A with eigenvalue zero, and we can extend $\{e_n \mid n \in \mathbb{N}\}$ to an orthonormal basis of $\mathcal{H}$ consisting of eigenvectors by adding an orthonormal basis of $\ker A$. □

A similar spectral theorem holds for compact, normal operators, which have orthogonal eigenvectors but possibly complex eigenvalues. A generalization also holds for bounded, self-adjoint or normal operators. In that case, however, the sum in (9.10) must be replaced by an integral of orthogonal projections with respect to an appropriate spectral measure that accounts for the possibility of a continuous spectrum. We will not discuss such generalizations in this book. Non-normal matrices on finite-dimensional linear spaces may be reduced to a Jordan canonical form, but the spectral theory of non-normal operators on infinite-dimensional spaces is more complicated.

We will discuss unbounded linear operators in the next chapter. The above theory may be used to study an unbounded operator whose inverse is compact or, more generally, an unbounded operator whose resolvent is compact, meaning that $(\lambda I - A)^{-1}$ is compact for some $\lambda \in \rho(A)$. For example, a regular Sturm-Liouville differential operator has a compact, self-adjoint resolvent, which explains why it has a complete orthonormal set of eigenvectors.

9.4 Compact operators

Before we can apply the spectral theorem for compact, self-adjoint operators, we have to check that an operator is compact. In this section, we discuss some ways to do this, and also give examples of compact operators.

The most direct way to prove that an operator A is compact is to verify the definition by showing that if E is a bounded subset of $\mathcal{H}$, then the set $A(E) = \{Ax \mid x \in E\}$ is precompact. In many examples, this can be done by using an appropriate condition for compactness, such as the Arzelà-Ascoli theorem or Rellich's theorem. The following theorem characterizes precompact sets in a general, separable Hilbert space.

Theorem 9.17 Let E be a subset of an infinite-dimensional, separable Hilbert space $\mathcal{H}$.

(a) If E is precompact, then for every orthonormal set $\{e_n \mid n \in \mathbb{N}\}$ and every $\epsilon > 0$, there is an N such that

$$\sum_{n=N+1}^{\infty} |\langle e_n, x\rangle|^2 < \epsilon \qquad \text{for all } x \in E. \tag{9.13}$$

(b) If E is bounded and there is an orthonormal basis $\{e_n\}$ of $\mathcal{H}$ with the property that for every $\epsilon > 0$ there is an N such that (9.13) holds, then E is precompact.

Proof. First, we prove (a). A precompact set is bounded, so it is sufficient to show that if E is bounded and (9.13) does not hold, then E is not precompact. If (9.13) does not hold, then there is an $\epsilon > 0$ such that for each N there is an $x_N \in E$ with

$$\sum_{n=N+1}^{\infty} |\langle e_n, x_N\rangle|^2 \geq \epsilon. \tag{9.14}$$

We construct a subsequence of (x_N) as follows. Let $N_1 = 1$, and pick N_2 such that

$$\sum_{n=N_2+1}^{\infty} |\langle e_n, x_{N_1}\rangle|^2 \leq \frac{\epsilon}{4}.$$

Given N_k, pick N_{k+1} such that

$$\sum_{n=N_{k+1}+1}^{\infty} |\langle e_n, x_{N_k}\rangle|^2 \leq \frac{\epsilon}{4}. \tag{9.15}$$

We can always do this because the sum

$$\sum_{n=1}^{\infty} |\langle e_n, x_N\rangle|^2 = \|x_N\|^2$$

converges by Parseval's identity.

For any $N \geq 1$, we define the orthogonal projection P_N by

$$P_N x = \sum_{n=1}^{N} \langle e_n, x \rangle e_n.$$

For $k > l$, we have

$$\begin{aligned}
\|x_{N_k} - x_{N_l}\|^2 &\geq \|(I - P_{N_k})(x_{N_k} - x_{N_l})\|^2 \\
&\geq [\|(I - P_{N_k}) x_{N_k}\| - \|(I - P_{N_k}) x_{N_l}\|]^2 \\
&= \left[\sqrt{\sum_{n=N_k+1}^{\infty} |\langle e_n, x_{N_k}\rangle|^2} - \sqrt{\sum_{n=N_k+1}^{\infty} |\langle e_n, x_{N_l}\rangle|^2}\right]^2 \\
&\geq \left[\epsilon^{1/2} - (\epsilon/4)^{1/2}\right]^2 \\
&\geq \frac{\epsilon}{4},
\end{aligned}$$

where we have used (9.14), (9.15), and the fact that $N_k \geq N_{l+1}$. It follows that (x_{N_k}) does not have any convergent subsequences, contradicting the hypothesis that E is precompact.

To prove part (b), suppose that $\{e_n \mid n \in \mathbb{N}\}$ is an orthonormal basis with the stated property, and let (x_n) be any sequence in E. We will use a diagonalization argument to construct a convergent subsequence, thus proving that E is precompact. Without loss of generality we may assume that $\|x\| \leq 1$ for all $x \in E$. We choose $n_1 = 1$. Then

$$\|(I - P_{n_1}) x_n\| \leq 1 \qquad \text{for all } n \in \mathbb{N}.$$

Since $P_{n_1} x_n$ is in the finite-dimensional Hilbert subspace spanned by $e_1, \ldots, e_{n_1}$ for each $n \in \mathbb{N}$, there is subsequence $(x_{1,k})$ of (x_n) such that $P_{n_1} x_{1,k}$ converges. Therefore, we can pick the subsequence such that (see Exercise 1.18)

$$\|P_{n_1}(x_{1,k} - x_{1,l})\|^2 \leq \frac{1}{k} \qquad \text{for } k \leq l.$$

Next, we choose n_2 such that

$$\|(I - P_{n_2}) x_{1,k}\|^2 \leq \frac{1}{2} \qquad \text{for all } k \in \mathbb{N}.$$

This is possible because of (9.13). We then pick a subsequence $x_{2,k}$ of $x_{1,k}$ such that $(P_{n_2} x_{2,k})$ is Cauchy and

$$\|P_{n_2}(x_{2,k} - x_{2,l})\|^2 \leq \frac{1}{k} \qquad \text{for all } k \leq l.$$

Continuing in this way, we choose n_l such that

$$\|(I - P_{n_l})\, x_{l-1,k}\|^2 \leq \frac{1}{l} \qquad \text{for all } k \in \mathbb{N},$$

and then pick a subsequence $(x_{l,k})$ of $(x_{l-1,k})$ such that $(P_{n_l} x_{l,k})$ satisfies

$$\|P_{n_l}\,(x_{l,k} - x_{l,j})\|^2 \leq \frac{1}{k} \qquad \text{for all } k \leq j.$$

The diagonal sequence $(x_{k,k})$ is Cauchy, since

$$\|x_{m,m} - x_{n,n}\|^2 = \|P_k(x_{m,m} - x_{n,n})\|^2 + \|(I - P_k)\, x_{m,m} - x_{n,n}\|^2 \leq \frac{2}{k}$$

for all $m, n \geq k$. □

Example 9.18 Let $\mathcal{H} = \ell^2(\mathbb{N})$. The *Hilbert cube*

$$C = \{(x_1, x_2, \ldots, x_n, \ldots) \mid |x_n| \leq 1/n\}$$

is closed and precompact. Hence C is a compact subset of $\mathcal{H}$.

Example 9.19 The diagonal operator $A : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ defined by

$$A(x_1, x_2, x_3, \ldots, x_n, \ldots) = (\lambda_1 x_1, \lambda_2 x_2, \ldots, \lambda_n x_n, \ldots), \tag{9.16}$$

where $\lambda_n \in \mathbb{C}$ is compact if and only if $\lambda_n \to 0$ as $n \to \infty$. Any compact, normal operator on a separable Hilbert space is unitarily equivalent to such a diagonal operator.

Proposition 5.43 implies that the uniform limit of compact operators is compact. An operator with finite rank is compact. Therefore, another way to prove that A is compact is to show that A is the limit of a uniformly convergent sequence of finite-rank operators. One such class of compact operators is the class of Hilbert-Schmidt operators.

Definition 9.20 A bounded linear operator A on a separable Hilbert space $\mathcal{H}$ is *Hilbert-Schmidt* if there is an orthonormal basis $\{e_n \mid n \in \mathbb{N}\}$ such that

$$\sum_{n=1}^{\infty} \|Ae_n\|^2 < \infty. \tag{9.17}$$

If A is a Hilbert-Schmidt operator, then

$$\|A\|_{HS} = \sqrt{\sum_{n=1}^{\infty} \|Ae_n\|^2} \tag{9.18}$$

is called the *Hilbert-Schmidt norm* of A.

One can show that the sum in (9.17) is finite in every orthonormal basis if it is finite in one orthonormal basis, and the norm (9.18) does not depend on the choice of basis.

Theorem 9.21 A Hilbert-Schmidt operator is compact.

Proof. Suppose that A is Hilbert-Schmidt and $\{e_n \mid n \in \mathbb{N}\}$ is an orthonormal basis. If P_N is the orthogonal projection onto the finite-dimensional space spanned by $\{e_1, \dots e_N\}$, then $P_N A$ is a finite-rank operator, and one can check that $P_N A \to A$ uniformly as $N \to \infty$. □

Example 9.22 The diagonal operator $A : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ defined in (9.16) is Hilbert-Schmidt if and only if

$$\sum_{n=1}^{\infty} |\lambda_n|^2 < \infty.$$

We say that A is a *trace class operator* if

$$\sum_{n=1}^{\infty} |\lambda_n| < \infty.$$

A trace class operator is Hilbert-Schmidt, and a Hilbert-Schmidt operator is compact.

Example 9.23 Let $\Omega \subset \mathbb{R}^n$. One can show that an integral operator K on $L^2(\Omega)$,

$$Kf(x) = \int_\Omega k(x,y) f(y)\, dy, \tag{9.19}$$

is Hilbert-Schmidt if and only if $k \in L^2(\Omega \times \Omega)$, meaning that

$$\int_{\Omega\times\Omega} |k(x,y)|^2 \, dxdy < \infty.$$

The Hilbert-Schmidt norm of K is

$$\|K\|_{HS} = \left(\int_{\Omega\times\Omega} |k(x,y)|^2 \, dxdy \right)^{1/2}.$$

If K is a self-adjoint, Hilbert-Schmidt operator then there is an orthonormal basis $\{\varphi_n \mid n \in \mathbb{N}\}$ of $L^2(\Omega)$ consisting of eigenvectors of K, such that

$$\int_\Omega k(x,y)\varphi_n(y)\, dy = \lambda_n \varphi_n(x).$$

Then

$$k(x,y) = \sum_{n=1}^{\infty} \lambda_n \varphi_n(x) \varphi_n(y),$$

where the series converges in $L^2(\Omega \times \Omega)$:

$$\lim_{N\to\infty} \int_{\Omega\times\Omega} \left| k(x,y) - \sum_{n=1}^{N} \lambda_n \varphi_n(x)\varphi_n(y) \right|^2 dxdy = 0.$$

For a proof, see, for example, Hochstadt [22].

Another way to characterize compact operators on a Hilbert space is in terms of weak convergence.

Theorem 9.24 A bounded linear operator on a Hilbert space is compact if and only if it maps weakly convergent sequences into strongly convergent sequences.

Proof. First, we show that a bounded operator $A : \mathcal{H} \to \mathcal{H}$ on a Hilbert space $\mathcal{H}$ maps weakly convergent sequences into weakly convergent sequences. If $x_n \rightharpoonup x$ as $n \to \infty$, then for every $z \in \mathcal{H}$ we have

$$\langle Ax_n - Ax, z\rangle = \langle x_n - x, A^* z\rangle \to 0 \qquad \text{as } n \to \infty.$$

Therefore, $Ax_n \rightharpoonup Ax$ as $n \to \infty$. Now suppose that A is compact, and $x_n \rightharpoonup x$. Since a weakly convergent sequence is bounded, the sequence (Ax_n) is contained in a compact subset of $\mathcal{H}$. Moreover, each strongly convergent subsequence is weakly convergent, so it converges to the same limit, namely Ax. It follows that the whole sequence converges strongly to Ax (see Exercise 1.27).

Conversely, suppose that A maps weakly convergent sequences into strongly convergent sequences, and E is a bounded set in $\mathcal{H}$. If (y_n) is a sequence in $A(E)$, then there is a sequence (x_n) in E such that $y_n = Ax_n$. By Theorem 8.45, the sequence (x_n) has a weakly convergent subsequence (x_{n_k}). The operator A maps this into a strongly convergent subsequence (y_{n_k}) of (y_n). Thus $A(E)$ is compact for any bounded set E, so A is compact. □

9.5 Functions of operators

The theory of functions of operators is called *functional calculus*. In this section, we describe some basic ideas of functional calculus in the special case of compact, self-adjoint operators.

If $q : \mathbb{C} \to \mathbb{C}$ is a polynomial function of degree d,

$$q(x) = \sum_{k=0}^{d} c_k x^k,$$

with coefficients $c_k \in \mathbb{C}$, then we define an operator-valued polynomial function

$q : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ in the obvious way as

$$q(A) = \sum_{k=0}^{d} c_k A^k. \tag{9.20}$$

There are several ways to define more general functions of a linear operator than polynomials. We have already seen that if A is a bounded operator and the function $f : \mathbb{C} \to \mathbb{C}$ is analytic at zero, with a Taylor series whose radius of convergence is strictly greater than $\|A\|$, then we may define $f(A)$ by a norm-convergent power series. For example e^A is defined for any bounded operator A, and $(I - A)^{-1}$ is defined in this way for any operator A with $r(A) < 1$.

An alternative approach is to use spectral theory to define a continuous function of a self-adjoint operator. First suppose that

$$A = \sum_{n=1}^{N} \lambda_n P_n$$

is a finite linear combination of orthogonal projections P_n with orthogonal ranges, and q is the polynomial function defined in (9.20). Since $\{P_n\}$ is an orthogonal family of projections, we have

$$P_n^k = P_n, \quad P_n P_m = 0 \qquad \text{for } n \neq m,\ k \geq 1.$$

It follows that $A^k = \sum_{n=1}^{N} \lambda_n^k P_n$ and

$$q(A) = \sum_{n=1}^{N} q(\lambda_n) P_n.$$

If A is a compact, self-adjoint operator with the spectral representation

$$A = \sum_{n=1}^{\infty} \lambda_n P_n, \tag{9.21}$$

then one can check that (see Exercise 9.18)

$$q(A) = \sum_{n=0}^{\infty} q(\lambda_n) P_n. \tag{9.22}$$

If $f : \sigma(A) \to \mathbb{C}$ is a continuous function, then a natural generalization of the expression in (9.22) for $q(A)$ is

$$f(A) = \sum_{n=1}^{\infty} f(\lambda_n) P_n. \tag{9.23}$$

This series converges strongly for any continuous f, and uniformly if in addition $f(0) = 0$ (see Exercise 9.19). An equivalent way to define $f(A)$ is to choose a sequence (q_n) of polynomials that converges uniformly to f on $\sigma(A)$, and define

$f(A)$ as the uniform limit of $q_n(A)$ (see Exercise 9.18). If f is real-valued, then the operator $f(A)$ is self-adjoint, and if f is complex-valued, then $f(A)$ is normal.

As a consequence of the spectral representation of A in (9.21) and $f(A)$ in (9.23), we have the following result.

Theorem 9.25 (Spectral mapping) If A is a compact, self-adjoint operator on a Hilbert space and $f : \sigma(A) \to \mathbb{C}$ is continuous, then

$$\sigma(f(A)) = f(\sigma(A)).$$

Here, $\sigma(f(A))$ is the spectrum of $f(A)$, and $f(\sigma(A))$ is the image of the spectrum of A under f,

$$f(\sigma(A)) = \{\mu \in \mathbb{C} \mid \mu = f(\lambda) \text{ for some } \lambda \in \sigma(A)\}.$$

A result of this kind is called a spectral mapping theorem. A spectral mapping theorem holds for bounded operators on a Hilbert space, and many unbounded operators, but there exist nonnormal, unbounded operators for which it is false (see Exercise 10.19). Thus, in general, unlike the finite-dimensional case, a knowledge of the spectrum of an unbounded operator is not sufficient to determine the spectrum of a function of the operator, and some knowledge of the operator's structure is also required.

Consider a linear evolution equation that can be written in the form

$$x_t = Ax, \qquad x(0) = x_0, \tag{9.24}$$

where A is a compact, self-adjoint linear operator on a Hilbert space $\mathcal{H}$. The solution is

$$x(t) = e^{At}x_0.$$

If x_0 is an eigenvector of A with eigenvalue λ, then $e^{At}x_0 = e^{\lambda t}x_0$. The solution decays exponentially if $\operatorname{Re}\lambda < 0$, and grows exponentially if $\operatorname{Re}\lambda > 0$. From the spectral mapping theorem, if the spectrum $\sigma(A)$ is contained in a left-half plane $\{\lambda \in \mathbb{C} \mid \operatorname{Re}\lambda \leq \omega\}$, then the spectrum of e^{At} is contained in the disc $\{\lambda \in \mathbb{C} \mid |\lambda| \leq e^{\omega t}\}$.

If $\{e_n \mid n \in \mathbb{N}\}$ is an orthonormal basis of eigenvectors of A, then we may write the solution as

$$x(t) = \sum_{n=1}^{\infty} e^{\lambda_n t}\langle e_n, x_0\rangle e_n.$$

If $\lambda_n \leq \omega$ for all n, it follows that

$$\|x(t)\| = \sqrt{\sum_{n=1}^{\infty} |e^{\lambda_n t}\langle e_n, x_0\rangle|^2} \leq e^{\omega t}\sqrt{\sum_{n=1}^{\infty} |\langle e_n, x_0\rangle|^2} = e^{\omega t}\|x_0\|.$$

When $\omega < 0$, any solution decays exponentially to 0 as $t \to \infty$. In that case, we say that the equilibrium solution $x(t) = 0$ is *globally asymptotically stable.*

9.6 Perturbation of eigenvalues

Suppose that $A(\epsilon)$ is a family of operators on a Hilbert space that depends on a real or complex parameter ϵ. If we know the spectrum of $A(0)$, then we can use perturbation theory to obtain information about the spectrum of $A(\epsilon)$ for small ϵ. In this section, we consider the simplest case, when $A(\epsilon)$ is a compact, self-adjoint operator depending on a real parameter ϵ.

Before doing this, we prove a preliminary result of independent interest: the Fredholm alternative for a compact, self-adjoint perturbation of the identity.

Theorem 9.26 Suppose that A is a compact, self-adjoint operator on a Hilbert space and $\lambda \in \mathbb{C}$ is nonzero. Then the equation

$$(A - \lambda I)\, x = y \tag{9.25}$$

has a solution if and only if $y \perp z$ for every solution z of the homogeneous equation

$$(A - \lambda I)\, z = 0.$$

The solution space of the homogeneous equation is finite-dimensional.

Proof. If $A : \mathcal{H} \to \mathcal{H}$ is compact and self-adjoint, then there is an orthonormal basis $\{e_n \mid n \in \mathbb{N}\}$ of $\mathcal{H}$ consisting of eigenvectors of A, with $Ae_n = \lambda_n e_n$ for $\lambda_n \in \mathbb{R}$. We expand x and y with respect to this basis as

$$x = \sum_{n=1}^{\infty} x_n e_n, \qquad y = \sum_{n=1}^{\infty} y_n e_n,$$

where $x_n = \langle e_n, x\rangle$ and $y_n = \langle e_n, y\rangle$. With respect to this basis, equation (9.25) has the diagonalized form

$$(\lambda_n - \lambda)\, x_n = y_n \qquad \text{for } n \in \mathbb{N}. \tag{9.26}$$

If $\lambda_n \neq \lambda$ for all n, then $\lambda_n - \lambda$ is bounded away from zero, since $\lambda \neq 0$ and there are no nonzero accumulation points of the eigenvalues of a compact operator. Hence, equation (9.25) is uniquely solvable for every $y \in \mathcal{H}$, with the solution

$$x = \sum_{n=1}^{\infty} \frac{\langle e_n, y\rangle}{\lambda_n - \lambda} e_n.$$

If $\lambda = \lambda_n$ for some n, then there is a finite-dimensional subspace of eigenvectors with the nonzero eigenvalue λ. Suppose the corresponding eigenvectors are

$\{e_{n_1}, e_{n_2}, \ldots, e_{n_k}\}$. Then we can solve (9.26) if and only if

$$y_{n_1} = y_{n_2} = \ldots = y_{n_k} = 0,$$

meaning that y is orthogonal to the kernel of $(A - \lambda I)$. □

Suppose that $A(\epsilon)$ is a compact, self-adjoint operator depending on a parameter $\epsilon \in \mathbb{R}$. We assume that A is a real-analytic function of ϵ at $\epsilon = 0$, meaning that it has a Taylor series expansion

$$A(\epsilon) = A_0 + \epsilon A_1 + \epsilon^2 A_2 + O(\epsilon^3)$$

that converges with respect to the operator norm in some interval $|\epsilon| < R$. The coefficient operators A_n are given by

$$A_0 = A(0), \quad A_1 = \dot{A}(\epsilon)\Big|_{\epsilon=0}, \quad A_2 = \frac{1}{2}\,\ddot{A}(\epsilon)\Big|_{\epsilon=0}, \ldots, \quad A_n = \frac{1}{n!}\,\frac{d^n}{d\epsilon^n} A(\epsilon)\Big|_{\epsilon=0}, \ldots$$

where the dot denotes a derivative with respect to ϵ.

We look for eigenvalues $\lambda(\epsilon)$ and eigenvectors $x(\epsilon)$ of $A(\epsilon)$ that satisfy

$$[A(\epsilon) - \lambda(\epsilon)]\, x(\epsilon) = 0. \tag{9.27}$$

It can be shown that the eigenvalues and eigenvectors of $A(\epsilon)$ have convergent Taylor series expansions

$$\lambda(\epsilon) = \lambda_0 + \epsilon\lambda_1 + \epsilon^2\lambda_2 + O(\epsilon^3), \tag{9.28}$$
$$x(\epsilon) = x_0 + \epsilon x_1 + \epsilon^2 x_2 + O(\epsilon^3), \tag{9.29}$$

where $\lambda_0 = \lambda(0)$, $\lambda_1 = \dot{\lambda}(0)$, and so on. We will not prove the convergence of these series here, but we will show how to compute the coefficients.

Setting $\epsilon = 0$ in (9.27), we obtain that

$$(A_0 - \lambda_0 I)\, x_0 = 0. \tag{9.30}$$

Thus, λ_0 is a nonzero eigenvalue of A_0 and x_0 is an eigenvector. For definiteness, we assume that λ_0 is a *simple eigenvalue* of A_0, meaning that it has multiplicity one, although eigenvalues of higher multiplicity can be treated in a similar way.

Differentiation of (9.27) with respect to ϵ implies that

$$(A - \lambda I)\,\dot{x} + \left(\dot{A} - \dot{\lambda} I\right) x = 0.$$

Setting ϵ equal to zero in this equation, we obtain that

$$(A_0 - \lambda_0 I)\, x_1 = \lambda_1 x_0 - A_1 x_0. \tag{9.31}$$

Since $\{x_0\}$ is a basis of $\ker(A_0 - \lambda_0 I)$, Theorem 9.26 implies that this equation is solvable for x_1 if and only if the right-hand side is orthogonal to x_0. It follows that

$$\lambda_1 = \frac{\langle x_0, A_1 x_0\rangle}{\|x_0\|^2}.$$

Continuing in this way, we differentiate equation (9.27) n times with respect to ϵ and set ϵ equal to zero in the result, which gives an equation of the form

$$(A_0 - \lambda_0 I)\, x_n = \lambda_n x_0 + f_{n-1}\left(x_0, \ldots, x_{n-1}, \lambda_0, \ldots, \lambda_{n-1}\right). \tag{9.32}$$

The Fredholm alternative implies that the right-hand side must be orthogonal to x_0. This condition determines λ_n, and we can then solve the equation for x_n. Thus, we can successively determine the coefficients in the expansions of $\lambda(\epsilon)$ and $x(\epsilon)$.

The solution of (9.32) for x_n includes an arbitrary multiple of x_0. This nonuniqueness is a consequence of the arbitrariness in the normalization of the eigenvector. If $c(\epsilon) = 1 + \epsilon c_1 + \epsilon^2 c_2 + O(\epsilon^3)$ is a scalar and $x(\epsilon) = x_0 + \epsilon x_1 + \epsilon^2 x_2 + O(\epsilon^3)$ is an eigenvector, then

$$c(\epsilon)x(\epsilon) = x_0 + \epsilon\left(x_1 + c_1 x_0\right) + \epsilon^2\left(x_2 + c_1 x_1 + c_2 x_0\right) + O(\epsilon^3)$$

is also an eigenvector. Each term in the expansion contains an arbitrary multiple of x_0.

An alternative way to derive the perturbation equations (9.30)–(9.32) is to use the Taylor series (9.28)–(9.29) in (9.27), expand, and equate coefficients of powers of ϵ in the result.

Example 9.27 Consider the eigenvalue problem

$$-u'' + V(x,\epsilon)u = \lambda u, \tag{9.33}$$

where $u \in L^2(\mathbb{R})$ and

$$V(x,\epsilon) = x^2 + \epsilon(x^4 - 4x^2)e^{-2x^2}. \tag{9.34}$$

In quantum mechanics, this problem corresponds to the determination of the energy levels of a slightly anharmonic oscillator. See Figure 9.1 for a graph of the potential for four different values of ϵ. For definiteness, we consider the energy of the ground state only, that is, the smallest eigenvalue, although the perturbation of other eigenvalues can be computed in exactly the same way.

The eigenvalue problem is of the form $Au = \lambda u$, where $A = A_0 + \epsilon A_1$ with

$$A = -\frac{d^2}{dx^2} + V, \quad A_0 = -\frac{d^2}{dx^2} + x^2, \quad A_1 = (x^4 - 4x^2)e^{-2x^2}.$$

From Exercise 6.14, the unperturbed ground state u_0 and the associated eigenvalue λ_0 of A_0 are given by

$$u_0(x) = e^{-x^2/2}, \qquad \lambda_0 = 1.$$

The operator A is unbounded. We will assume that the perturbed operator has a ground state close to that of the unperturbed operator, and apply the above expansion without discussing the validity of the method in this case. See the book by Kato [26] for a comprehensive discussion.

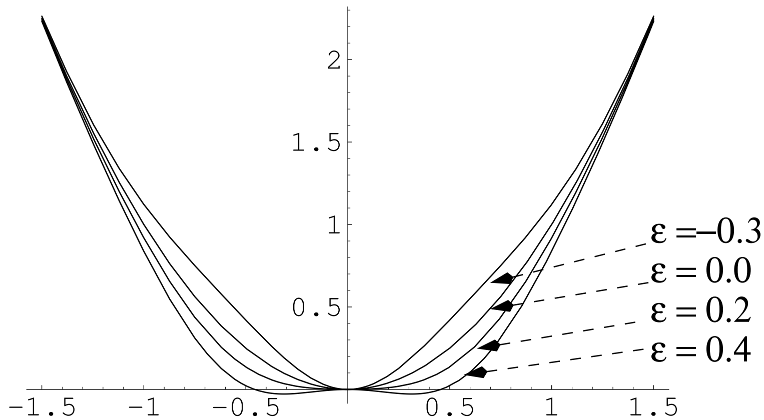

Fig. 9.1 The perturbed harmonic potential $V(x, \epsilon)$, defined in (9.34), for ϵ=-0.3, 0, 0.2, and 0.4.

The perturbed eigenvalue has the expansion $\lambda = 1 + \epsilon\lambda_1 + O(\epsilon^2)$, where

$$\lambda_1 = \frac{\langle u_0, (x^4 - 4x^2)e^{-2x^2}u_0\rangle}{\|u_0\|^2} = \frac{\int_{-\infty}^{\infty}(x^4 - 4x^2)e^{-3x^2}\,dx}{\int_{-\infty}^{\infty} e^{-x^2}\,dx}. \qquad (9.35)$$

The expression in (9.35) for λ_1 may be evaluated in the following way. First, note that

$$\frac{\int_{-\infty}^{\infty} x^n e^{-3x^2}\,dx}{\int_{-\infty}^{\infty} e^{-x^2}\,dx} = \frac{1}{3^{(n+1)/2}}\frac{\int_{-\infty}^{\infty} x^n e^{-x^2}\,dx}{\int_{-\infty}^{\infty} e^{-x^2}\,dx},$$

for $n \geq 0$. We need to compute this ratio of integrals for $n = 2$ and $n = 4$. Let

$$J(a) = \int_{-\infty}^{\infty} e^{-x^2+2ax}\,dx. \qquad (9.36)$$

Differentiating this expression n times with respect to a, we obtain that

$$J^{(n)}(a) = 2^n \int_{-\infty}^{\infty} x^n e^{-x^2+2ax}\,dx.$$

Hence, setting $a = 0$, we have

$$\frac{\int_{-\infty}^{\infty} x^n e^{-x^2}\,dx}{\int_{-\infty}^{\infty} e^{-x^2}\,dx} = \frac{1}{2^n}\frac{J^{(n)}(0)}{J(0)}.$$

To evaluate the right-hand side of this equation, we complete the square in the exponent of the integrand in (9.36) and change the integration variable from $x \mapsto x - a$. This gives

$$J(a) = e^{a^2} J(0).$$

It follows that

$$\frac{\int_{-\infty}^{\infty} x^n e^{-x^2}\,dx}{\int_{-\infty}^{\infty} e^{-x^2}\,dx} = \frac{1}{2^n} \left.\frac{d^n e^{a^2}}{da^n}\right|_{a=0}.$$

In particular, for $n = 2, 4$, we compute

$$\frac{d^2 e^{a^2}}{da^2} = \left(4a^2 + 2\right) e^{a^2}, \quad \frac{d^4 e^{a^2}}{da^4} = \left(16a^4 + 48a^2 + 12\right) e^{a^2}.$$

From these computations it follows that $\lambda_1 = -7/(24\sqrt{3})$, and

$$\lambda = 1 - \frac{7}{24\sqrt{3}}\epsilon + O(\epsilon^2).$$

If $\epsilon > 0$, corresponding to an oscillator that becomes "softer" for small amplitude oscillations, then the ground state energy decreases, while if $\epsilon < 0$, corresponding to an oscillator that becomes "stiffer" for small amplitude oscillations, then the ground state energy increases.

9.7 References

For additional discussion of the spectra of bounded and compact, normal operators, see Naylor and Sell [40]. The terminology of the classification of the spectrum is not entirely uniform (see Reed and Simon [45] for a further discussion). With the definitions we use here, the spectrum of a bounded operator is the disjoint union of its point, continuous, and residual spectrums. For an introduction to complex analysis and a proof of Liouville's theorem, see [36]. See Kato [26] for the perturbation theory of spectra.

9.8 Exercises

Exercise 9.1 Prove that $\rho(A^*) = \overline{\rho(A)}$, where $\overline{\rho(A)}$ is the set $\{\lambda \in \mathbb{C} \mid \overline{\lambda} \in \rho(A)\}$.

Exercise 9.2 If λ is an eigenvalue of A, then $\overline{\lambda}$ is in the spectrum of A^*. What can you say about the type of spectrum $\overline{\lambda}$ belongs to?

Exercise 9.3 Suppose that A is a bounded linear operator on a Hilbert space and $\lambda, \mu \in \rho(A)$. Prove that the resolvent R_λ of A satisfies the *resolvent equation*

$$R_\lambda - R_\mu = (\mu - \lambda)\, R_\lambda R_\mu.$$

Exercise 9.4 Prove that the spectrum of an orthogonal projection P is either $\{0\}$, in which case $P = 0$, or $\{1\}$, in which case $P = I$, or else $\{0, 1\}$.

Exercise 9.5 Let A be a bounded, nonnegative operator on a complex Hilbert space. Prove that $\sigma(A) \subset [0, \infty)$.

Exercise 9.6 Let G be a multiplication operator on $L^2(\mathbb{R})$ defined by

$$Gf(x) = g(x)f(x),$$

where g is continuous and bounded. Prove that G is a bounded linear operator on $L^2(\mathbb{R})$ and that its spectrum is given by

$$\sigma(G) = \overline{\{g(x) \mid x \in \mathbb{R}\}}.$$

Can an operator of this form have eigenvalues?

Exercise 9.7 Let $K : L^2([0,1]) \to L^2([0,1])$ be the integral operator defined by

$$Kf(x) = \int_0^x f(y)\, dy.$$

(a) Find the adjoint operator K^*.
(b) Show that $\|K\| = 2/\pi$.
(c) Show that the spectral radius of K is equal to zero.
(d) Show that 0 belongs to the continuous spectrum of K.

Exercise 9.8 Define the right shift operator S on $\ell^2(\mathbb{Z})$ by

$$S(x)_k = x_{k-1} \qquad \text{for all } k \in \mathbb{Z},$$

where $x = (x_k)_{k=-\infty}^{\infty}$ is in $\ell^2(\mathbb{Z})$. Prove the following facts.

(a) The point spectrum of S is empty.
(b) $\operatorname{ran}(\lambda I - S) = \ell^2(\mathbb{Z})$ for every $\lambda \in \mathbb{C}$ with $|\lambda| > 1$.
(c) $\operatorname{ran}(\lambda I - S) = \ell^2(\mathbb{Z})$ for every $\lambda \in \mathbb{C}$ with $|\lambda| < 1$.
(d) The spectrum of S consists of the unit circle $\{\lambda \in \mathbb{C} \mid |\lambda| = 1\}$ and is purely continuous.

Exercise 9.9 Define the *discrete Laplacian operator* Δ on $\ell^2(\mathbb{Z})$ by

$$(\Delta x)_k = x_{k-1} - 2x_k + x_{k+1}, \tag{9.37}$$

where $x = (x_k)_{k=-\infty}^{\infty}$. Show that $\Delta = S + S^* - 2I$. Prove that the spectrum of Δ is entirely continuous and consists of the interval $[-4, 0]$.
HINT: Consider $x_k = e^{ik\xi}$ where $-\pi \leq \xi \leq \pi$. Finite difference schemes for the numerical solution of differential equations may be written in terms of shift operators, and a study of their spectrum is useful in the stability analysis of finite difference schemes (see Strikwerder [53]).

Exercise 9.10 Define the right shift operator S on $\ell^2(\mathbb{N})$ by

$$S((x_1, x_2, x_3, \ldots)) = (0, x_1, x_2, \ldots), \tag{9.38}$$

and the left shift operator T on $\ell^2(\mathbb{N})$ by

$$T((x_1, x_2, x_3, \ldots)) = (x_2, x_3, x_4, \ldots). \tag{9.39}$$

Prove the following.

(a) The resolvent set of S is the exterior of the unit disc $\{\lambda \in \mathbb{C} \mid |\lambda| > 1\}$.
(b) Every $\lambda \in \mathbb{C}$ with $|\lambda| = 1$ belongs to the continuous spectrum of S.
(c) $\operatorname{ran}(\lambda I - S)$ is not dense in ℓ^2 for every $\lambda \in \mathbb{C}$ with $|\lambda| < 1$, meaning that the interior of the unit disc is contained in the residual spectrum of S.
(d) The resolvent set of T consists of all $\lambda \in \mathbb{C}$ such that $|\lambda| > 1$.
(e) The continuous spectrum of T is the unit circle.
(f) The point spectrum of T is the interior of the unit disc.
(g) The residual spectrum of T is empty.

Exercise 9.11 A complex number λ belongs to the *approximate spectrum* of a bounded linear operator $A : \mathcal{H} \to \mathcal{H}$ on a Hilbert space $\mathcal{H}$ if there is a sequence (x_n) of vectors in $\mathcal{H}$ such that $\|x_n\| = 1$ and $(A - \lambda I)x_n \to 0$ as $n \to \infty$. Prove that the approximate spectrum is contained in the spectrum, and contains the point and continuous spectrum. Give an example to show that a point in the residual spectrum need not belong to the approximate spectrum.

Exercise 9.12 Let $\mathcal{H}$ be a separable Hilbert space with an orthonormal basis $\{e_n\}$, and $A \in \mathcal{B}(\mathcal{H})$ such that

$$\sum_n \|Ae_n\|^2 < \infty.$$

(a) Prove that the Hilbert-Schmidt norm defined in (9.18) is independent of the basis. That is, show that for any other orthonormal basis $\{f_n\}$ one has

$$\sum_n \|Af_n\|^2 = \sum_n \|Ae_n\|^2.$$

(b) Prove that

$$\|A\|_{HS} = \|A^*\|_{HS}.$$

Exercise 9.13 Suppose that $L : \mathbb{R} \to \mathcal{B}(\mathcal{H})$ and $A : \mathbb{R} \to \mathcal{B}(\mathcal{H})$ are smooth, operator-valued functions of $t \in \mathbb{R}$, where $L(t)$ is self-adjoint and $A(t)$ is skew-adjoint. If $L(t)$ satisfies the ODE

$$\dot{L} = [L, A], \tag{9.40}$$

show that

$$L(t) = U^*(t)L(0)U(t),$$

where $\dot{U} = UA$ and $U(0) = I$. Show that $U(t)$ is unitary, and deduce that the eigenvalues of $L(t)$ are independent of t.

An equation that can be written in the form (9.40) for suitable operators $L(t)$ and $A(t)$ is said to be *completely integrable* because it possesses a large number of conserved quantities, namely, the eigenvalues of L. The pair of operators $\{L, A\}$ is called a *Lax pair* for the equation.

Exercise 9.14 Show that the $n \times n$, tridiagonal matrices

$$L = \begin{pmatrix} b_1 & a_1 & 0 & \dots & 0 \\ a_1 & b_2 & a_2 & \dots & 0 \\ 0 & a_2 & b_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & b_n \end{pmatrix}, \quad A = \begin{pmatrix} 0 & a_1 & 0 & \dots & 0 \\ -a_1 & 0 & a_2 & \dots & 0 \\ 0 & -a_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

form a Lax pair for the *Toda lattice equations*

$$\begin{aligned} \dot{a}_k &= a_k\left(b_{k+1} - b_k\right) \qquad \text{for } k = 1, \dots, n-1, \\ \dot{b}_k &= 2\left(a_k^2 - a_{k-1}^2\right) \qquad \text{for } k = 1, \dots, n, \end{aligned}$$

where $a_0 = a_n = 0$ and $a_k > 0$. Write out the equations for $n = 2$, and determine explicitly their conserved quantities.

Exercise 9.15 Show that the KdV equation for $u(x,t)$,

$$u_t = 6uu_x - u_{xxx},$$

can be written in the form (9.40), where L and A are the following differential operators acting on smooth functions of x:

$$L = -\partial_x^2 + u, \qquad A = -4\partial_x^3 + 3u\partial_x + 3\partial_x u.$$

Exercise 9.16 Prove that (9.20) and (9.22) define the same operator $q(A)$.

Exercise 9.17 Let A be a compact, self-adjoint operator, on an infinite-dimensional separable Hilbert space, and $q : \sigma(A) \to \mathbb{C}$ a continuous function.

(a) Prove that the series in (9.22) converges in norm if and only if $q(0) = 0$.
(b) For an arbitrary value $q(0)$, prove that the series in (9.22) converges strongly.

Exercise 9.18 Suppose that A is a compact self-adjoint operator. Let $f \in C(\sigma(A))$, and consider $f(A)$ defined by (9.23). Prove that

$$\|f(A)\| = \sup\{|f(\lambda_n)| \mid n \in \mathbb{N}\}.$$

Let (q_N) be a sequence of polynomials of degree N, converging uniformly to f on $\sigma(A)$. The existence of such a sequence is a consequence of the Weierstrass approximation theorem. Prove that $(q_N(A))$ converges in norm, and that its limit equals $f(A)$ as defined in (9.23).

Exercise 9.19 Let A be a compact selfadjoint linear operator. Prove that the series in (9.23) is convergent in the strong operator topology for any $f \in C(\sigma(A))$, and that it converges uniformly if in addition $f(0) = 0$.

Exercise 9.20 Let A be a self-adjoint compact operator on a Hilbert space $\mathcal{H}$, and let $f : \sigma(A) \to \mathbb{C}$ be a continuous function. When is $f(A)$ compact?

Exercise 9.21 Consider the evolution equation $x_t = Ax$, where A is a bounded operator on a Hilbert space such that

$$\operatorname{Re}\langle x, Ax\rangle \leq 2\alpha\|x\|^2,$$

for some $\alpha \in \mathbb{R}$. By taking the inner product of the evolution equation with x, derive the energy estimate

$$\|x(t)\| \leq e^{\alpha t}\|x(0)\|.$$

Compare this result with that of the spectral method for self-adjoint and non-self-adjoint operators A.

Exercise 9.22 Suppose that A is a compact, nonnegative linear operator on a Hilbert space. Prove that there is a unique compact, nonnegative linear operator B such that $B^2 = A$. Thus, $B = A^{1/2}$ is the square root of A.

Exercise 9.23 Consider the eigenvalue problem

$$\int_{-\infty}^{\infty} e^{-|x|-|y|}u(y)\,dy + \epsilon x u(x) = \lambda u(x), \qquad -\infty < x < \infty,$$

where $\epsilon \in \mathbb{R}$. Show that if $\epsilon = 0$, then the spectrum consists purely of eigenvalues, and $\lambda = 1$ is a simple eigenvalue with eigenfunction $u(x) = e^{-|x|}$. Show that a formal perturbation expansion with respect to ϵ as $\epsilon \to 0$ gives

$$\begin{aligned}\lambda &= 1 + \frac{1}{2}\epsilon^2 + \dots,\\ u(x) &= e^{-|x|}\left(1 + \epsilon x + \epsilon^2\left(x^2 - \frac{3}{4}\right) + \dots\right).\end{aligned}$$

Show, however, that there are no eigenfunctions $u \in L^2(\mathbb{R})$ when $\epsilon \neq 0$. (It is possible to show that then the spectrum is purely continuous.)